PB81-166704

Theoretical and Empirical Evaluation of Multipoint Rollback as Applied to Smelters in the Hayden, Arizona Area

SRI International
Menlo Park, CA

Prepared for

Environmental Protection Agency
San Francisco, CA

May 79

U.S. Department of Commerce
National Technical Information Service
NTIS.

PB81-166706

THEORETICAL AND EMPIRICAL EVALUATION OF MULTIPOINT ROLLBACK AS APPLIED TO SMELTERS IN THE HAYDEN, ARIZONA AREA

Final Report

May 1979

Environmental Protection Agency Region 9

JUN 13 1979

LIBRARY

By: Richard H. Thuillier, Senior Research Meteorologist
Atmospheric Science Center

Harold L. Javitz, Senior Statistician
Urban and Social Systems Division

Prepared for:

U.S. Environmental Protection Agency Region IX
Air and Hazardous Materials Division
215 Fremont Street
San Francisco, California 94105

Attn: Dr. Norman Grib

and for

U.S. Environmental Protection Agency
401 M Street, Southwest
Washington, D.C. 20460

Attn: Mr. John R. Busik (EN-341)
Mr. Steve Hitte (EN-341)
Ms. Marian Bernd (PM-214-C)

Contract No. 68-01-4137
Task No. 13

SRI Project 5828-13

Approved:

Warren B. Johnson, Director
Atmospheric Science Center

Earle D. Jones, Executive Director
Advanced Development Division

SRI International

SRI International

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA 22161

50272-101

REPORT DOCUMENTATION PAGE	1. REPORT NO.	2.	3. Recipient's Accession No. PB81 166704

4. Title and Subtitle
THEORETICAL AND EMPIRICAL EVALUATION OF MULTIPOINT ROLLBACK AS APPLIED TO SMELTERS IN THE HAYDEN, ARIZONA AREA

5. Report Date May 1979

6.

7. Author(s)
Richard H. Thuillier & Harold L. Javitz

8. Performing Organization Rept. No. 5828-13

9. Performing Organization Name and Address
SRI International
333 Ravenswood Ave.
Menlo Park, CA 94025

10. Project/Task/Work Unit No. 13

11. Contract(C) or Grant(G) No.
(C) 68-01-4137
(G)

12. Sponsoring Organization Name and Address
U.S. Environmental Protection Agency
215 Fremont St.
San Francisco, CA 94105

13. Type of Report & Period Covered

14.

15. Supplementary Notes

16. Abstract (Limit: 200 words)

In 1978, EPA promulgated a final rule for the control of sulfur oxide emissions at existing nonferrous smelters in Arizona. Emission limits, determined from a 1973-74 data base, were set by proportionally rolling back the 24-hourly average smelter emissions estimated by sulfur balance or other methods for the day on which the maximum three-hour-average ambient concentration of sulfur dioxide was observed. In response to the EPA rule, an alternative approach to calculating copper smelter emission limits based on a statistical model was proposed by Arizona. EPA contracted with SRI International for an evaluation of both rules. The analysis suggests that emission limits imposed by the current EPA rule provide for a greater confidence in achieving ambient air quality standards than do the limits suggested in the report by Arizona; and that the feasibility of implementing the multipoint rollback method at a smelter to produce a specific enforceable distribution of emissions has yet to be demonstrated.

17. Document Analysis a. Descriptors
Copper
Arizona
Enforcement
Air pollution control

b. Identifiers/Open-Ended Terms
Hayden AZ
Copper smelters
Multipoint rollback
Emission control

c. COSATI Field/Group

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA 22161

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages 26
	20. Security Class (This Page)	22. Price

C23447

CONTENTS

I INTRODUCTION 1

II CONCLUSIONS 3

III DISCUSSION 4

A. Theoretical Evaluation of Multipoint Rollback . . . 5

B. Empirical Evaluation of the Multipoint and Single-Point Rollback Techniques 7

C. Implementation and Enforcement 11

APPENDIX

A REVIEW OF THE MULTIPOINT ROLLBACK METHOD A-1

I INTRODUCTION

On 4 January 1978, the EPA promulgated a final rule for the control of sulfur oxide emissions at existing nonferrous smelters in the State of Arizona, imposing a maximum emission limitation for each smelter as a six-hour running mean. Emission limits, determined from a 1973-74 data base, were set by proportionally rolling back the 24-hourly average smelter emissions estimated by sulfur balance or other methods for the day on which the maximum three-hour-average ambient concentration of sulfur dioxide was observed. The EPA rule is to be enforced by continuous in-stack monitoring of sulfur dioxide concentration and flow rate.

In response to the EPA rule, the Kennecott Company has proposed an alternative approach to calculating copper smelter emission limits based on a statistical model. The Arizona Department of Health Services commissioned the University of Arizona to review the proposed approach and this review resulted in yet another approach for calculating emission limits. This new approach is known as multipoint rollback and is based upon rollback of the entire distribution of emissions at the smelter. Multipoint rollback is described in a paper by the University of Arizona which favors this approach over the proportional rollback approach of EPA. EPA's Region IX Office, which has ultimate responsibility for ensuring that the National Ambient Air Quality Standards are achieved, has been requested to consider the multipoint rollback approach as a replacement for the current EPA rule.

EPA's Region IX Office has determined the need for technical assistance in evaluating Arizona's proposed method, and has contracted with SRI International for an objective evaluation of both the EPA rule and the statistical method as formulated by the University of Arizona. Specifically, the EPA needs to know:

- Whether the statistical method is theoretically sound.
- Whether the available data base is adequate for a confident application of the statistical approach.
- Whether the statistical method will improve confidence in achieving air-quality standards over that associated with the EPA rule.
- Whether the method is enforceable and feasible to implement.

To achieve the goals of EPA's Region IX, SRI reviewed in detail the multipoint rollback method for control of emissions from copper smelters in the State of Arizona. Documents reviewed consisted of "Sulfur Dioxide Emission Limits for Hayden Copper Smelters" by D. B. Morris and L. S. Salmon of Kennecott Company; "A Review of Multi-Point

Rollback: A Method for Establishing Ultimate Emissions Limitations for Copper Smelters in Arizona" by J. L. Moyers and T. W. Peterson of the University of Arizona; "Technical Support Document: Final Regulations for Control of Sulfur Dioxide Emissions from Arizona Copper Smelters" by EPA Region IX; and "First and Second Annual Reports on Arizona Copper Smelter Pollution Control Technology" by C. Billings of the Arizona Department of Health Services. Personal Interviews were also conducted with Dr. Norman Grib and Mr. Larry Bowerman of EPA, Region IX, and with Dr. Thomas Peterson and Dr. Jarvis Moyers of the University of Arizona.

II CONCLUSIONS

Based on the arguments in this Technical Memorandum, we must conclude that while the multipoint rollback approach has merit from a theoretical standpoint, its application based on the 1976 Hayden area data base alone may not be valid. Specifically, our conclusions are as follows:

1) The statistical method is theoretically sound when all the enabling assumptions are satisfied.

2) Certain key assumptions and statistical requirements are not satisfied with regard to smelter operations in the Hayden area. Specifically:

   - The requirement of proportionality between ambient concentrations and stack emissions, under a given set of meteorological conditions, is probably not valid due primarily to the influence of fugitive emissions which do not exit through the stack.
   - The requirement of independence between emissions and meteorological conditions is probably violated because of the use at the Asarco smelter of supplemental control which, by definition, relates emissions to meteorological conditions.
   - The assumption of attainment of air quality standards when the expected frequency of violation has been reduced to once-per-year is highly uncertain when based on a single year of data. Even with statistically ideal data, it can be demonstrated that a data set covering a single year will provide no more than a 75% confidence in achievement of the air quality standards.

3) Because the above deficiencies are reflected in the 1976 data from the Hayden area, this data base should not be used in its present form as the basis for setting emission limits for Hayden area or other Arizona smelters. The possibility of accounting for the deficiencies through statistical manipulation of the data does exist but further study will be needed to ascertain the feasibility of such an approach. Ideally, a data set should be obtained in which enabling assumptions are satisfied if multipoint rollback is to be applied.

4) Analysis of the 1976 data base, granted its deficiencies, suggests that emission limits imposed by the current EPA rule,

being more stringent, provide for a greater confidence in achieving ambient air quality standards than do the limits suggested in the report by the University of Arizona. A more meaningful comparison of methods must await the availability of an improved data base.

5) The feasibility of implementing the multipoint rollback method at a smelter to produce a specific enforceable distribution of emissions has yet to be demonstrated.

III DISCUSSION

A. Theoretical Evaluation of Multipoint Rollback

The validity of the multipoint rollback approach depends on the correctness of the mathematical formulation, the validity of assumptions on which the mathematics is based, and the reasonableness of the conclusions drawn from the mathematical exercise.

The general conclusion from the multipoint rollback formulation is that, given distributions of smelter emissions and ambient sulfur dioxide concentrations for the same period of time at a specified site, a modified distribution of emissions can be determined statistically that will guarantee the achievement of air quality standards, at the site in question, with a given degree of confidence. An infinite variety of emissions distributions might provide the necessary guarantee, but the one specifically recommended by the University of Arizona would roll back each value in the current distribution of smelter emissions by a given factor. The rollback factor proposed by the University of Arizona for the Hayden, Arizona area is the ratio of the air quality standard concentration to the maximum concentration observed at the Montgomery Ranch monitoring site for the appropriate averaging time.

In scrutinizing the method as outlined by the University of Arizona in their review paper, we found that the general conclusion is theoretically valid, given certain enabling assumptions. We found it necessary, however, to reformulate the mathematical derivation to provide a more rigorous treatment (see Appendix). The assumptions that must be satisfied are presented in the following paragraphs along with a discussion of their validity.

Assumption I--Ambient concentration averaged over a given time interval is proportional to total smelter emissions averaged over a concurrent or lagging time period of equal length. The constant of proportionality is determined by meteorological conditions so that a given emission will always yield the same ambient concentration for a given set of meteorological conditions.

The assumption of the proportionality of emissions and concentrations based on meteorological conditions is both the multipoint and single-point (EPA) methods. Except for complications that might arise from chemical species transformation, it is a physically realistic assumption. The multipoint rollback approach, however, as presently formulated, assumes the proportionality for a single source and discrete time intervals on a one-to-one basis. The method overlooks the possible relationship of a single air-quality value to emissions in more than one time interval and from more than one source point. Since the complex

terrain setting for the Hayden smelters results in complex flows involving possible recirculation of emitted pollutants, more than one time interval of emissions could possibly contribute to a single air-quality value. Since smelter emissions do not emanate from a single point (that is, both stack emissions and low-level fugitive emissions at two spatially separated smelters are involved), similar meteorological conditions may lead to considerably dissimilar ambient SO_2 concentrations, depending on the distribution of emissions among the various source points.

For the reasons just stated--multiple contributing time intervals and source points--the assumption of proportionality of total smelter emissions to ambient concentrations under similar meteorological conditions is likely not satisfied for the environment of the Hayden smelters.

Assumption II--The meteorological conditions and emissions are independent. The meaning of this assumption is that a given emission is equally likely to coincide with any one of the many possible sets of meteorological conditions occurring in the Hayden area. This is a reasonable assumption for an operation that proceeds independently of meteorological conditions.

In the case of the Hayden smelters, a supplemental control system (SCS) was in operation during 1976. Conversations with Mr. J. Peckenpaw at Kennecott smelter and Mr. R. Guptill at Asarco smelter revealed that the SCS was operational during 1976 and that operations at Asarco were keyed to meteorological factors. Specifically, the number of roasters and converters feeding the Asarco stack was reduced in adverse meteorological situations to avoid violations of air quality standards. It is quite probable, therefore, that the 1976 data set contains fewer and lower values in excess of the standard than would be the case if SCS were not in use. It must therefore be concluded that the emission-concentration data set that served as a demonstration data base for the multipoint rollback approach, is "contaminated" by SCS. This leads us to conclude that computation of the rollback factor as proposed by the University of Arizona, based on the 1976 data, might not provide for sufficient rollback.

Assumption III--Emissions during a given time period and the lag between the time of emission and the resulting measured concentration are independent. This assumption is reasonable, and we have found no evidence causing us to doubt its validity in general. For a data set contaminated by SCS, the assumption could be invalidated if, for instance, wind direction was used as a control criterion.

Assumption IV--If the probability of a violation equals 1/N, where N is the number of discrete time periods in a 1-year period, attainment of the air quality standard is assured. This assumption presupposes that the standard is attained if, on the average, no more than one

violation per year occurs. This statistical form of the standard obviously allows more than one violation in some years, balanced by a lack of violations in other years. It must be noted that this statistical definition is at variance with the current definition of the SO_2 standard, which specifies no more than one violation in any given year.

The assumption also presupposes that any data set used to estimate the probability of violation is of sufficient size to provide an acceptable degree of confidence in the guarantee of attainment. Given independent sets of nonoverlapping 3-hour averages of emissions and SO_2 concentrations, and given that all 3-hour values are independent of one another, it can be calculated that approximately 8000 3-hour periods (about 2-3/4 years) would be required for a guarantee of standard attainment, as currently defined, with 95% confidence. A single year of data satisfying the same statistical conditions listed above would provide only a 75% chance of attaining the standard in a given year. In reality, the confidence must be less than the ideal values stated above because the lengthy nature of some adverse meteorological episodes implies that some successive 3-hour averages are dependent (i.e., one influences another) during such episodes. The impact of such dependence is to increase the size of the data set required for a given degree of confidence in attaining the standard. These statements apply to both the single-point and the multipoint approach.

From the arguments put forth above, we must conclude that the assumptions required for confident implementation of either the multipoint or single-point rollback method are probably violated at the Hayden site, especially with use of the 1976 data base. Furthermore, the technique of deriving the rollback factor from a single value in the annual distribution of all air quality values allows for a large margin of error, especially in a data set influenced by supplemental control. In the next section the two methods will be compared empirically using the 1976 Hayden area data set.

B. Empirical Evaluation of the Multipoint and Single-Point Rollback Techniques

Because of the many flaws in the assumptions underlying the theoretical validity of the multipoint rollback and EPA approaches as applied to the Hayden area smelter operations, an empirical examination of the effectiveness of the models is warranted. We chose to conduct this examination using the 1976 data set on which the multipoint rollback approach was demonstrated in the University of Arizona report.

1. The Data Set

The data set in question consisted of concurrent half-hour averaged values of smelter stack emissions in units of pounds per hour, and ambient sulfur dioxide (SO_2) concentrations in units of micrograms per

cubic meter. Emissions data were available on magnetic tape for the Kennecott and Asarco smelters individually and for the two smelters combined. Concentration data were available for the Montgomery Ranch and Crescent Ranch monitoring sites in the vicinity of the Hayden smelters. For our analysis of the data set, we chose to use total emissions for the two Hayden smelters and ambient concentrations at the Montgomery Ranch monitoring site. This choice is consistent with the data used by the University of Arizona in their own demonstration of multipoint rollback. Also, maximum concentrations observed at Montgomery Ranch during 1976 exceeded those at Crescent Ranch by almost a factor of 2.

An examination of the data set prior to analysis revealed a considerable amount of zero drift in the concentration values as evidenced by the presence of negative concentration values in the data set. Table 1 provides a distribution of positive and negative concentration values as well as an indication of the number of missing half-hour values stratified by time of day. Missing values accounted for about 12% and negative values for about 30% of total half-hour periods available. In utilizing the data for our analysis we first computed hourly averages, ignoring the missing data. That is, two half-hour values were averaged; a single half-hour value was used as the hourly average when the other value was missing, and the hourly average was treated as missing when both half-hour values were missing. Running 3-hour, 6-hour, and 24-hour means were computed from the hourly values only when all hours during the averaging time had data available.

2. Empirical Examination of Method Effectiveness

The effectiveness of the single-point and multipoint rollback approaches and peculiarities of the data set are graphically presented in Figures 1 and 2. Figure 1 is a scatterplot of 3-hour running mean SO_2 concentration versus the 6-hour running mean of total (Kennecott plus Asarco) smelter emissions. The 3-hour concentration period is in all cases the last 3 hours of the 6-hour period used for emissions, since emissions at one time are not logically associated with concentrations at an earlier time. Of 86 violations appearing on the scatter plot, 35 are nonoverlapping in time.

From the figure it may be seen that most violations of the 3-hour SO_2 standard of 1300 $\mu g/m^3$ occur with relatively low values of the 6-hour emission rate. Furthermore, high-concentration values are conspicuously absent above the 50,000-lb/hr emission rate. This situation, which is strongly counterintuitive, suggests the use of supplemental control during the 1976 period. Exceedances of the 1300-$\mu g/m^3$ level primarily at low emission rates may be due to residual stack emissions under SCS (a minimum of one converter and one roaster feeding the stack were required at Asarco by company policy), or quite possibly to fugitive emissions that were not directly affected by SCS.

Vertical dashed lines on the figure indicate, from left to right, the 6-hour emission limit currently imposed by EPA regulations based on

Table 1

SO_2 CONCENTRATION FREQUENCY DISTRIBUTION FOR HAYDEN, ARIZONA DURING 1976

HALF/CLAS1:	.GT.	0	-25	-50	-75	-100	-150	-200	-250	-300	-350	-400	-450	.LT.	MSG.
HOUR/CLAS2:	0	-25	-50	-75	-100	-150	-200	-250	-300	-350	-400	-450	-500	-500	DATA
.5	245	59	17	4	0	0	1	0	0	0	1	0	0	0	39
1.0	256	53	12	4	0	0	1	0	0	0	0	0	0	0	38
1.5	260	49	11	6	0	0	0	1	0	0	0	0	0	0	39
2.0	266	51	11	5	0	0	0	1	0	0	0	0	0	0	32
2.5	267	48	11	4	0	0	1	1	0	0	0	0	0	0	34
3.0	273	39	14	4	0	1	0	1	0	0	0	0	0	0	34
3.5	266	43	16	4	1	0	0	1	0	0	0	0	0	0	35
4.0	267	41	16	5	1	1	0	1	0	0	0	0	0	0	34
4.5	266	43	13	5	0	1	0	1	0	0	0	0	0	0	37
5.0	272	36	16	5	0	0	0	2	0	0	0	0	0	0	35
5.5	260	27	18	4	0	0	1	1	0	0	0	0	0	0	35
6.0	266	36	15	5	0	1	0	2	0	0	0	0	0	0	37
6.5	270	39	13	5	0	1	0	2	0	0	0	0	0	0	36
7.0	273	36	14	6	1	0	0	2	0	0	0	0	0	0	34
7.5	276	33	12	5	0	0	0	1	1	0	0	0	0	0	36
8.0	261	33	12	1	1	0	0	1	1	0	0	0	0	0	06
8.5	264	45	9	6	1	0	0	1	1	0	0	0	0	0	39
9.0	257	47	10	6	1	0	0	1	0	0	0	0	0	0	42
9.5	261	52	11	7	0	0	0	1	2	0	0	0	0	0	32
10.0	256	55	11	6	0	0	0	0	2	0	0	0	0	0	32
10.5	257	56	13	5	1	0	0	0	2	1	0	0	0	0	29
11.0	250	61	16	5	0	0	0	0	2	1	0	0	0	0	31
11.5	236	78	11	6	0	0	0	1	1	0	0	0	0	1	32
12.0	224	83	16	7	1	0	0	0	1	0	0	0	0	0	34
12.5	206	79	16	5	1	0	0	0	1	0	1	0	0	0	56
13.0	206	90	26	9	0	0	0	0	1	1	0	1	0	0	32
13.5	175	91	29	10	2	0	0	0	1	0	1	1	0	0	56
14.0	152	96	21	11	2	0	0	0	0	1	0	2	0	0	79
14.5	139	102	32	10	3	0	1	1	0	0	0	0	1	0	77
15.0	140	109	26	11	3	1	0	0	0	0	0	0	1	0	73
15.5	150	130	33	14	2	1	0	0	0	0	0	0	0	1	35
16.0	137	132	30	13	3	1	0	0	0	0	0	0	0	1	49
16.5	130	116	23	10	1	1	0	0	0	0	0	0	0	1	64
17.0	139	139	36	12	1	1	0	0	0	0	0	0	3	1	37
17.5	140	126	36	12	2	1	0	0	0	0	0	0	0	1	44
18.0	130	133	36	12	2	1	0	0	0	0	0	0	0	1	49
18.5	129	133	34	14	1	2	0	0	0	0	0	0	0	1	52
19.0	140	130	34	15	1	2	0	0	0	0	0	0	0	1	43
19.5	152	123	33	14	0	2	0	0	0	0	0	0	0	1	41
20.0	156	117	33	12	1	2	0	0	0	0	0	0	0	1	44
20.5	164	115	30	11	1	2	0	0	0	0	0	0	0	1	42
21.0	169	104	31	12	1	2	0	0	0	0	0	0	0	1	46
21.5	162	91	36	10	1	2	0	0	0	0	0	0	0	1	43
22.0	169	95	30	9	0	3	0	0	0	0	1	0	0	0	40
22.5	196	90	31	6	0	2	1	0	0	0	0	0	0	0	39
23.0	213	75	26	10	0	2	0	0	0	0	0	0	0	0	38
23.5	223	72	26	7	0	0	2	0	0	0	0	0	0	0	36
24.0	242	65	16	4	0	0	1	0	0	0	1	0	0	0	35
TOTALS:	10275	3704	1034	379	36	33	9	23	16	4	5	4	2	14	2032

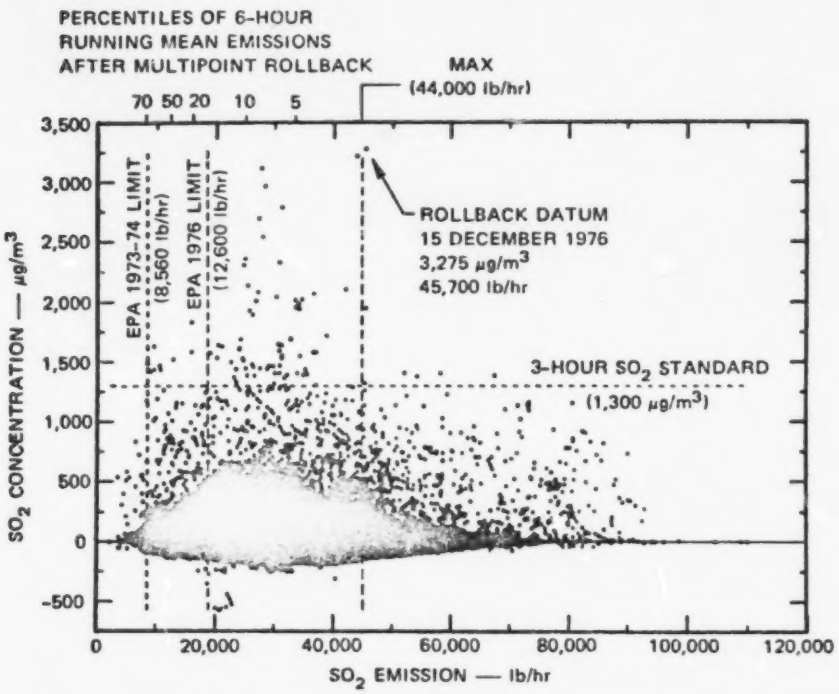

FIGURE 1 SCATTER PLOT OF 3-HOUR RUNNING MEAN SO_2 CONCENTRATIONS AT THE MONTGOMERY RANCH SITE VERSUS TOTAL 6-HOUR RUNNING MEAN SO_2 EMISSIONS FROM KENNECOTT AND ASARCO SMELTERS

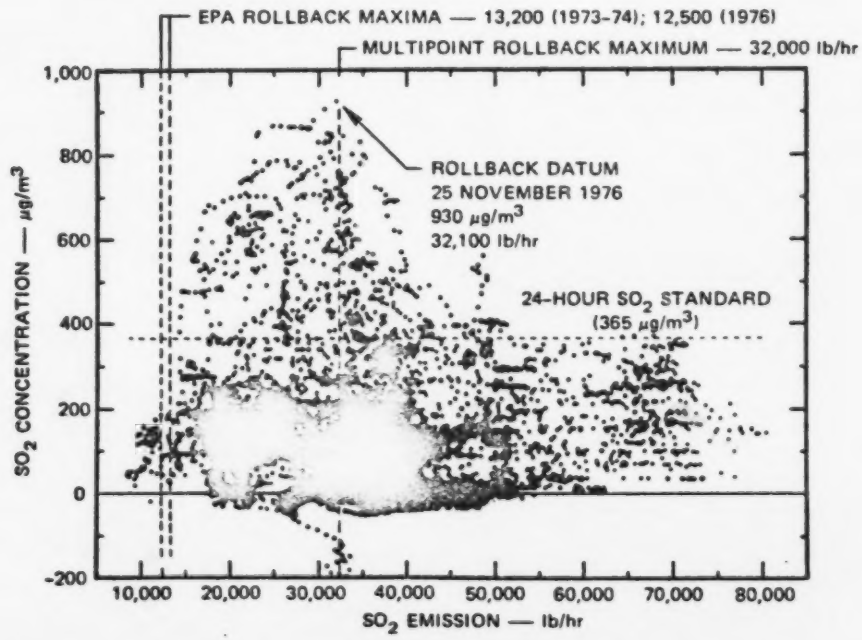

FIGURE 2 SCATTER PLOT OF 24-HOUR RUNNING MEAN SO_2 CONCENTRATIONS AT THE MONTGOMERY RANCH SITE VERSUS TOTAL 24-HOUR RUNNING MEAN SO_2 EMISSIONS FROM KENNECOTT AND ASARCO SMELTERS

1973-74 data, the 6-hour emission limit that would have been imposed by EPA had 1976 data been used and the 6-hour emission maximum provided by the multipoint rollback technique. Along the top of the figure is indicated a scale that represents the percentile 6-hour emission rate values allowed by multipoint rollback. The figure indicates that the current EPA limit of 8560 lb/hr is almost precisely the value at which 1976 violations cease to exist. In view, however, of the small sample of emissions below the limit, violations below the limit are not beyond the realm of possibility. With the multipoint limits based on the data set, 70% of allowable emissions would be at levels associated with violations in 1976. This suggests that given the conditions in 1976, multipoint rollback would allow many more violations than would the EPA's single-point approach whether the EPA rollback was based on 1973-74 or on 1976 data. The data also suggest that the stringency of the EPA method is warranted if the air quality standard is to be achieved at the Montgomery Ranch site. If in fact the violations observed at lower emission rates are due to fugitive emissions, apparently the EPA's 8560 lb/hr emission limit is stringent enough to account even for such fugitive emissions.

Figure 2 is a scatter plot of running mean concentration and emission values each for the same 24-hour period. The data contain many overlapping values, and only 32 of these represent individual 24-hour periods, midnight to midnight. The lack of violations at the higher emission rates is particularly striking in the 24-hour data. Vertical dashed lines from left to right represent the respective 24-hour emission limits given by the EPA single-point approach using 1973-74 and 1976 data and the maximum 24-hour emission rate allowable under the multipoint rollback scheme. The 24-hour limits are less stringent because concentrations in excess of the limits of Figure 1 would obviously occur. The 3-hour standard is therefore the controlling standard. As in the case of the 3-hour standard, multipoint rollback limits would potentially allow many more violations than the single-point method, given the conditions of 1976.

C. Implementation and Enforcement

Since the EPA single-point rollback approach specifies merely an upper limit to emissions, implementation of the method is straightforward, albeit drastic and/or costly to the smelter industry. Large reductions in smelter feed rate or extensive control of emissions from all phases of smelter operations would be required. Enforcement by in-stack monitoring is possible, although such monitoring currently fails to account adequately for fugitive emissions.

In the case of the multipoint rollback scheme, it would seem that a somewhat sophisticated control approach might be required to produce a specific enforceable distribution of emissions. It is not clear at this point just how this would be accomplished. As far as enforcement is concerned, however, the same in-stack monitoring approach would apply

except that emission limits would be checked for recurrence frequency rather than for a single instance of exceedance. However, an enforceable regulation implementing such limits has not yet been drafted. In any event, the success of any monitoring program would depend on all smelter emissions passing through the stack being monitored.

Appendix

A REVIEW OF THE MULTIPOINT ROLLBACK METHOD*

The multipoint rollback method has been suggested as a method for establishing ultimate emissions limitations for copper smelters in Arizona. In this appendix we discuss the statistical properties of this method. In particular we shall be concerned with the assumptions that underlie the method and that must be satisfied before the method could guarantee compliance with air quality standards.

Our first task is to specify a model relating pollutant emissions to pollutant concentrations, such that the model leads (by statistical reasoning) to the development of the multipoint rollback method. To the extent to which the model is correct, the multipoint rollback method could guarantee compliance with air quality standards. The second task is to evaluate the extent to which the model is correct, or is an oversimplification of the actual processes relating emissions to concentrations. Thus, in the first task we establish the modeling assumptions that allow the statistical development of the multipoint rollback method and in the second task we evaluate the validity of those assumptions.

1. Modeling Assumptions

Some of the modeling assumptions we are presenting in this section do not appear in the University of Arizona review of the multipoint rollback method. These assumptions are consequences of certain modeling implications of the statistical statements that appear in the review. After discussing these implications with Dr. Peterson of the University of Arizona, SRI was able to construct a model that rigorously defines and clarifies these modeling concepts, and that allowed us to correct some technical errors in the statistical equations. We emphasize that the main benefit of clearly defining a model that allows the development of the multipoint rollback method is that we can then understand the

*This appendix was prepared by H. L. Javitz.

assumptions that must be satisfied in order for the multipoint rollback method to statistically guarantee compliance with air quality standards.*

In order to specify the model, we first establish some notation:

(1) Let t be an index that denotes time intervals. These time intervals may be of any length (e.g., 15 min, 1 hr, 3 hr, 24 hr). Thus, when t = 1, we are discussing the first time interval, when t = 2, we are discussing the second time interval, etc. We require that these time intervals be nonoverlapping.

(2) Let E_t be the average emissions in the t-th time period and let C_t be the ambient pollutant concentrations in the t-th time period. Thus, E_5 is the average emission in the 5th time period and C_{17} is the ambient concentration in the 17th time period.

A model is now proposed to relate the sequence of emissions $(E_1, E_2, E_3, \ldots, E_{50}, \ldots)$ to the sequence of concentrations $(C_1, C_2, C_3, \ldots, C_{50}, \ldots)$. The model states that at a particular time t (e.g., at the 250th time interval) the measured concentration C_t is caused by an emission E_{t-L} emitted at time t - L. Thus, if L = 4, the model is stating that it took four time periods for the emission to reach the air quality monitoring site. We have chosen the letter L to denote "lag." The lag need not be constant over time. For example, at t = 250 the lag might be L = 4, while at t = 117 the lag might be L = 6. The lag may also be equal to zero. Since the lag changes with time, we will provide the lag with a subscript t. Thus L_t is the lag between the time t at which the concentration was monitored and the time of the causal emission (e.g., $L_{250} = 4$ and $L_{117} = 6$).

The model also specifies the relationship between the magnitude of the concentration C_t at time t and the causal emission E_{t-L_t}. The emission E_{t-L_t} has been subjected to certain meteorological conditions from the time of emission to the time t when the resultant concentration is measured. The effect of these meteorological conditions may be summarized in a meteorological (dispersion) parameter

$$D_t = \frac{C_t}{E_{t-L_t}} .$$

*It may be possible to apply the multipoint rollback method under a model which has fewer assumptions than the one presented here. It is clearly desirable to do so.

Thus the model states that the concentration C_t observed at time t is the product of the meteorological parameter D_t and the causal emission E_{t-L_t}. Thus the first assumption is:

$$(I) \quad C_t = D_t * E_{t-L_t} .$$

Using assumption I we may write the probabilistic equation giving us the probability that C_t is more than some arbitrary concentration value χ (e.g. $\chi = 57\ \mu g/m^3$) as follows:

$$\text{Prob}\{C_t > \chi\} = \sum_{\ell=0}^{\infty} \text{Prob}\{L_t = \ell;\ D_t * E_{t-\ell} > \chi\} . \quad \ldots \quad (1)$$

We pause to comment on a technical point. The probability Prob $\{C_t > \chi\}$ is undoubtedly a function of t. For example, during periods, when production has stopped, it is unlikely that C_t is large. Similarly, during months that usually have poor dispersion it is likely that C_t will be large. In order to avoid difficulties of this sort we shall specify that the value of t in the previous equation has been randomly selected from among all possible values during the year. Thus, Prob $\{C_t > \chi\}$ is the probability that on a randomly selected time period during the year the measured concentration would exceed χ.

In order to proceed with the statistical derivation we must define the (conditional) probability that $D_t * E_{t-\ell}$ is more than χ when $L_t = \ell$. This (conditional) probability is denoted

$$\text{Prob}\{D_t * E_{t-\ell} > \chi \mid L_t = \ell\}$$

and is defined as equal to

$$\frac{\text{Prob}\{L_t = \ell;\ D_t * E_{t-\ell} > \chi\}}{\text{Prob}\{L_t = \ell\}} .$$

Using this definition we may rewrite our probabilistic equation as

$$\text{Prob}\{C_t > \chi\} = \sum_{\ell=0}^{\infty} \left[\text{Prob}\{L_t = \ell\} * \text{Prob}\{D_t * E_{t-\ell} > \chi \mid L_t = \ell\}\right] . \quad (2)$$

The next step in the statistical derivation is to rewrite $\text{Prob}\{D_t * E_{t-\ell} > \chi \mid L_t = \ell\}$ using integrals. Now both D_t and $E_{t-\ell}$ are continuous variables. Let their joint probabilistic density function when $L_t = \ell$ be denoted by

$$p(D_t, E_{t-\ell} \mid L_t = \ell) .$$

The probability that $D_t * E_{t-\ell} > \chi$ when $L_t = \ell$ may be obtained by integrating this joint probabilistic density over the region of values for D_t and $E_{t-\ell}$ for which $D_t * E_{t-\ell} > \chi$. That is,

$$\begin{aligned} &\text{Prob}\{D_t * E_{t-\ell} > \chi \mid L_t = \ell\} \\ &= \iint_{D_t * E_{t-\ell} > \chi} p(D_t, E_{t-\ell} \mid L_t = \ell)\, dE_{t-\ell} dD_t \\ &= \int_0^{\infty} \left[\int_{\chi/D_t}^{\infty} p(D_t, E_{t-\ell} \mid L_t = \ell)\, dE_{t-\ell}\right] dD_t . \quad (3) \end{aligned}$$

In order to simplify the above equation the model assumes that

(II) For any given lag time $L_t = \ell$, the meteorological conditions D_t and the emissions $E_{t-\ell}$ are independent.

Using this assumption we may express the joint density of D_t and $E_{t-\ell}$ when $L_t = \ell$ as the product of the density of D_t when $L_t = \ell$, and the density of $E_{t-\ell}$ when $L_t = \ell$. That is,

$$p(D_t, E_{t-\ell} \mid L_t = \ell) = p(D_t \mid L_t = \ell)\, p(E_{t-\ell} \mid L_t = \ell) . \quad (4)$$

Since SCS control is not allowed, we may also assume:

(III) E_{t-L_t} and L_t are independent.

Using this assumption we may write

$$p(D_t, E_{t-\ell} \mid L_t = \ell) = p(D_t \mid L_t = \ell)p(E_{t-\ell}) \ .$$

Combining this result with the previous integral equation we may write

$$\text{Prob } \{D_t * E_{t-\ell} > \chi \mid L_t = \ell\}$$

$$= \int_0^\infty \left[\int_{\chi/D_t}^\infty p(D_t \mid L_t = \ell)\ p(E_{t-\ell})\ dE_{t-\ell} \right] dD_t$$

$$= \int_0^\infty p(D_t \mid L_t = \ell) \left[\int_{\chi/D_t}^\infty p(E_{t-\ell})\ dE_{t-\ell} \right] dD_t$$

$$= \int_0^\infty p(D_t \mid L_t = \ell)\ \text{Prob } \{E_{t-\ell} > \frac{\chi}{D_t}\}\ dD_t \ . \qquad (5)$$

Combining this last equation with Eq. (2), we obtain the long-sought-after result:

$$\text{Prob } \{C_t > \chi\}$$

$$= \sum_{\ell=0}^{\infty} \text{Prob } \{L_t = \ell\} \left[\int_0^\infty p(D_t \mid L_t = \ell)\ \text{Prob } \{E_{t-\ell} > \frac{\chi}{D_t}\}\ dD_t \right]. \qquad (6)$$

We desire that emissions be controlled in such a way that the probability of the concentration C_t exceeding the applicable standard C_{std} at a given time t be very small. Let this small probability be denoted "eps." Thus, we desire

$$\text{Prob } \{C_t > C_{std}\} = \text{eps} \ .$$

The reviewers at the University of Arizona have suggested that eps be set equal to 1/2920, if the time intervals are three hours in length,

since there are 2920 three-hour time intervals in a year. If we follow this suggestion, then we desire that emissions be controlled in such a way that the upper (1 - 1/2920) percentile of the probability distribution of C_t be equal to C_{std}.

We now ask the question "What is the actual upper (1 - 1/2920)th percentile of the probability distribution of C_t?" We do not know this quantity but we shall denote it by C_{eps}. Later on we shall use the data to estimate C_{eps}. With current emissions we have (by definition)

$$\text{Prob } \{C_t > C_{eps}\} = \text{eps} .$$

What can be done to the current distribution of emission so that we obtain the desired result Prob $\{C_t > C_{std}\}$ = eps ? One method is to institute control measures so that the resultant distribution of emissions is the same as the distribution of

$$\frac{C_{std}}{C_{eps}} * E_{t-\ell} .$$

To demonstrate the truth of this assertion, suppose that these control measures have been taken. By Eq. (6),

$$\text{Prob } \{C_t > C_{std} \text{ after control measures have been taken}\}$$

$$= \sum_{\ell=0}^{\infty} \text{Prob } \{L_t = \ell\} * \left[\int_0^{\infty} p(D_t \mid L_t = \ell) * \right. \tag{7}$$

$$\left. \text{Prob } \{E_{t-\ell} > \frac{C_{std}}{D_t} \text{ after control measures}\} * dD_t \right] .$$

By the design of the control measures,

$$\text{Prob}\{E_{t-\ell} > \frac{C_{std}}{D_t} \text{ after control measures}\}$$

$$= \text{Prob}\{\frac{C_{std}}{C_{eps}} * E_{t-\ell} > \frac{C_{std}}{D_t} \text{ before control}\}$$

$$= \text{Prob}\{E_{t-\ell} > \frac{C_{eps}}{D_t} \text{ before control}\} . \qquad (8)$$

Substituting this result into the previous equation, we obtain:

$$\text{Prob}\{C_t > C_{std} \text{ after control measures have been taken}\}$$

$$= \sum_{\ell=0}^{\infty} \text{Prob}\{L_t = \ell\} \left[\int_0^{\infty} p(D_t \mid L_t = \ell) * \right.$$

$$\left. \text{Prob}\{E_{t-\ell} > \frac{C_{eps}}{D_t} \text{ before control}\}\, dD_t \right]$$

$$= \text{Prob}\{C_t > C_{eps} \text{ before control}\} \qquad (9)$$

$$= \text{eps} .$$

We have now demonstrated that with multipoint rollback the probability that the concentration C_t will exeeed the C_{std} is only eps. To implement this rollback we need to estimate percentile C_{eps}. If eps = 1/2920, one natural estimate of C_{eps} is C_{max}, the largest observed concentration during the year. Using this estimate we are required to roll back emissions so that the "new" emissions have the same distribution as does C_{std}/C_{max} times the "old" emissions.

At this point in the statistical development we are able to write that with the multipoint rollback,

$$\text{Prob}\{C_t > C_{std}\} \doteq \frac{1}{2920} .$$

The air quality standards, however, state that C_{std} is not to be exceeded more than one time in any year. Thus, we really seek to know

$$\text{Prob}\{C_t > C_{std} \text{ more than once for } t = 1,2,3, \ldots, 2920\} .$$

The first probability statement is a statement about a single concentration at a random time t. The second probability statement is a statement about an entire year of concentration observations. Without further assumptions concerning the joint probability density of $\{C_1, C_2, C_3, \ldots\}$, it is not possible to determine the exact magnitude of the probability of no more than one violation in any year. For example, it is possible to specify joint probability densities* such that

$$\text{Prob}\{C_t > C_{std} \text{ more than once for } t = 1,2, \ldots, 2920\} = 1/2 \text{ or } 0$$

while

$$\text{Prob}\{C_t > C_{std}\} = \frac{1}{2920} .$$

One of the easiest assumptions to make would be that the $\{C_t : 1 \leq t \leq 2920\}$ are independent, with equal probabilities of violating the standard. Under this asusmption we can compute

$$\text{Prob}\{C_t > C_{std} \text{ more than once for } t = 1,2, \ldots, 2920\} = 0.266 .$$

Thus, there would be over a 25% probability that the standard would be violated more than once in any year. In order to be 95% confident that during one year the standard will not be violated more than once, it would be necessary to compute the rollback fraction using a C_{max} value obtained from approximately 8000 observations (about 2-3/4 years of observations).

---

*To achieve the first result (a probability of 1/2), suppose that

$$\text{Prob}\{C_{100} > C_{std}, C_{200} > C_{std}, \text{ other } C_i < C_{std}\} = 1/2$$

$$\text{Prob}\{\text{all } C_i < C_{std} \text{ for } i = 1,2, \ldots, 2920\} = 1/2$$

and for the second result (a probability of zero), suppose that

$$\text{Prob}\{C_{100} > C_{std}, \text{ other } C_i < C_{std}\} = 1 .$$

As a result of the above discussion we must add another assumption to the model in order to conclude that the multipoint rollback method would guarantee compliance with the air quality standards:

(IV) The joint probability density of $\{C_1, C_2 C_3, \ldots\}$ guarantees that if Prob $\{C_t > C_{std}\} = \frac{1}{2920}$, then Prob $\{C_t > C_{std}$ more than once for $t = 1,2, \ldots, 2920\} = 0.05$.

2. Validity of the Assumptions

The model assumes that concentrations at time t are the result of emissions during a single, earlier time $t - \ell$. This assumption may be satisfactory when dispersion conditions are good. When dispersion conditions are poor, however, the concentrations measured at a given time may have been caused by emissions at many previous times. That is, the emissions may accumulate in the valley. To the extent to which the accumulation of emissions during periods of poor dispersion results in high concentrations, the modeling assumption of the multipoint rollback method is overly optimistic. In this same vein we should comment that the EPA single-point rollback at least limits the amount of emissions per time period. It is likely that the highest emission rates naturally group together during the periods when substantial quantities of high sulfur ore are going through certain processing phases. Under the multipoint rollback method, the emissions might be higher for a number of time periods in such a processing phase than that allowed under the EPA single-point rollback method. If this processing phase coincided with a period of poor dispersion, then the accumulation of very large emissions might easily result in a concentration above the standard.

The model also assumes that D_t and $E_{t-\ell}$ are independent. As long as SCS controls are not in effect, this assumption appears to be reasonable. Similarly, we see no reason to take issue with the assumption that E_{t-L_t} and L_t are independent.

The last assumption concerning the joint probabilistic density of $\{C_1, C_2, C_3, \ldots\}$ is crucial. Without an assumption of this nature we cannot bridge the gap between a probability statement concerning the concentration at one randomly selected time and a probability statement concerning 2920 consecutive time intervals. As we previously state, the probability statement

$$P\{C_t > C_{std}\} \doteq \frac{1}{2920}$$

is consistent with having the probability of more than one concentration above C_{std} be as large as 1/2 or as small as 0. Under the assumption of independence and equal probability of violation we found that there was about a 26% chance of having two or more concentrations above C_{std} yearly. (Since the concentrations appear to be serially dependent, we would not place too much trust in the 26% figure). To the best of our knowledge, no arguments have been developed that would support assumption IV.